DOCUMENT RESUME

ED 151 421 TM 006 959

AUTHOR	McKenna, Bernard
TITLE	Task Force and Other Reports. Interim Report of the Task Force on Testing.
INSTITUTION	National Education Association, Washington, D.C.
PUB DATE	Jul 73
NOTE	22p.; Paper presented at the Annual Representative Assembly of the National Education Association (52nd, Portland, Oregon, July 3-6, 1973); For related document, see TM 006 646
EDRS PRICE	MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS	Bibliographies; Educational Testing; Evaluation Methods; *Instructional Improvement; *Standardized Tests; Student Testing; *Teacher Associations; Teacher Evaluation; *Testing Problems; *Testing Programs
IDENTIFIERS	*National Education Association; *NEA Task Force on Testing

ABSTRACT

The National Education Association's (NEA) resolutions concerning standardized tests, teacher evaluation and subjective ratings, national testing and assessment, student rights, and the improvement of instruction are presented in this interim report of the Task Force on Testing. The NEA strongly encourages the elimination of group standardized intelligence, aptitude, and achievement tests to assess student potential or achievement until completion of a critical appraisal, review, and revision of current testing programs. The Task Force recommendations for immediate action, for further study on testing by the National Education Association, and for other assessment-related issues are presented, and a bibliography is appended. (Author/MV)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

Task Force and Other Reports

Presented to the Fifty-Second Representative Assembly of the National Education Association

July 3-6, 1973
Portland, Oregon

ED151421

TM006 959

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Frances Quinto
NEA

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

Report of the Task Force on Practitioner Involvement
Interim Report of the Asian American Task Force
✓Interim Report of the Task Force on Testing
Legislative Committee Recommendations
Report of the Challenge Committee
Report of the Committee on Accountability

FINDINGS AND RECOMMENDATIONS

INTERIM REPORT

NEA TASK FORCE ON TESTING

Adopted by the Task Force, May 29, 1973

Prepared for the Task Force by Bernard McKenna

TASK FORCE ON TESTING

Chairman

Charles J. Sanders, Junior High Counselor
President-Elect, Maine Teachers Association
Millinocket Junior High School
Millinocket, Maine 04462
Business: (207) 723-9488
Home: (207) 723-4412

Members

Mrs. Jean Blachford, Classroom Teacher
201 Exeter Street
Highland Park, New Jersey 08904
Business: (201) 247-2600, Ext. 211
Home: (201) 572-3611

Lupe Castillo, Classroom Teacher
157 Bertita Street
San Francisco, California 94112
Business: (415) 261-8516
Home: (415) 585-6132

Mrs. Dorothy Lee Collins, Counselor
1217 Delaware Street
San Antonio, Texas 78210
Business: (512) 433-2361, Ext. 236
Home: (512) 532-4497

Robert L. Harcourt, Guidance Director
5937 Schoolwood Road
Indianapolis, Indiana 46224
Business: (317) 241-9285
Home: (317) 291-1042

Members (continued)

Ms. Jane B. Mace
Secondary Reading Specialist
2443 North Weil
Milwaukee, Wisconsin 53212
Business: (414) 562-5012
Home: (414) 964-1538

Mrs. Golda Walbert, Classroom Teacher
313 Leslie Avenue
Glasgow, Kentucky 42141
Home: (502) 651-3886

Ezekial Williams
Gambrell Hall
Benedict College
Columbia, South Carolina 29204
Business: (803) 779-4930

Staff Contact

Bernard McKenna
Instruction & Professional Development

Staff Consultants

Boyd Bosma, Teacher Rights
Alton B. Sheridan, Research

NEA RESOLUTIONS AND NEW BUSINESS ITEMS ON TESTING

72-44. Standardized Tests

The National Education Association strongly encourages the elimination of group standardized intelligence, aptitude, and achievement tests to assess student potential or achievement until completion of a critical appraisal, review, and revision of current testing programs.

NEA New Business Items, 1972

Testing

This Representative Assembly directs the National Education Association to immediately call a national moratorium on standardized testing and at the same time set up a task force on standardized testing to research and make its findings available to the 1975 Representative Assembly for further action. (Item 28)

The NEA shall establish a task force to deal with the numerous and complex problems communicated to it under the general heading of testing. This task force shall report its findings and proposals for further action at the 1973 Representative Assembly. (Item 51)

OTHER SUPPORTING RESOLUTIONS

C-6. Evaluation and Subjective Ratings

The National Education Association believes that it is a major responsibility of educators to participate in the evaluation of the quality of their services. To enable educators to meet this responsibility more effectively, the Association calls for continued research and experimentation to develop means of objective evaluation of the performance of all educators, including identification of (a) factors that determine professional competence; (b) factors that determine the effectiveness of competent professionals; (c) methods of evaluating effective professional service; and (d) methods of recognizing effective professional service through self-realization, personal status, and salary.

The Association also believes that evaluations should be conducted for the purpose of improvement of performance and quality of instruction offered to pupils, based upon written criteria and following procedures mutually developed by and acceptable to the teacher association, the administration and the governing board.

The Association insists that the evaluation program must recognize the rights of the educator who is evaluated. These include the right to:

a. Information concerning the evaluation procedure of the school district or institution.
b. Open evaluation without subterfuge and advance notice of evaluation visits with discussion of the teacher's goals and methods.
c. Evaluation at least in part by peers skilled in the teacher's professional or subject area.
d. Consultation in timely fashion after a formal evaluation visit and receipt of and opportunity to acknowledge in writing any formal evaluation report prior to placement in a personnel file.
e. Evaluation reports which assess strengths, note progress, indicate remaining deficiencies and suggest specific measures the teacher can take to overcome indicated deficiencies.
f. Participation in a professional development program including such activities as appropriate counseling and supportive services, released time for in-service work,

and opportunity to observe or seek and give assistance to other teachers in classroom settings other than one's own.

g. Review of any material considered derogatory prior to placement in the individual's personnel file and submission of a written answer attached to the item in the file.

h. Supervision which is constructive, provides an opportunity to correct deficiencies, takes into account the variety of learning and teaching environmental factors, and emphasizes career development of the professional educator.

The Association believes that examinations such as the National Teacher Examination must not be used as a condition of employment or a method for evaluating educators in service for purposes such as salary, tenure, retention, or promotion. (69, 70, 72)

72-13. National Testing and Assessment

The National Education Association notes that the first report of the National Assessment of Educational Progress on writing, citizenship and science has been issued.

The Association will continue to resist any attempt to transform assessment results into a national testing program that would seek to measure all students or school systems by a single standard, and thereby impose upon them a single program rather than providing opportunities for multiple programs and objectives.

72-8. Student Rights

The National Education Association believes that basic student rights include: the right to free inquiry and expression; the right to due process; the right to freedom of association; the right to freedom of peaceful assembly and petition; the right to participate in the governance of the school, college, and university; the right to freedom from discrimination; and the right to equal educational opportunity.

C-10. Improvement of Instruction

The National Education Association believes that a prime responsibility of professional associations is to stimulate significant improvements in the quality of instruction. Much of the responsibility to make educational changes should lie with the teachers through their influence and involvement in democratic decision making in and out of the school.

The Association supports the principle of involving its National Affiliates, Associated Organizations, and Departments in efforts to improve instruction in our schools.

The Association urges local affiliates to involve members and those affected in the development and implementation of programs for instructional improvement, curriculum development, and individualization of instruction relevant to the needs of the students.

The Association recommends that professional educators enter into active collaboration with research and development specialists, both in regional educational laboratories and in industry, to promote technology's potential contribution to education by guiding the development of technology in the most educationally sound directions. It encourages school systems to establish learning materials centers.

The Association further recommends that the profession, in cooperation with other interested groups, establish standards for educational materials, and insist that publishers and producers use the services of a competent educational institution or facility to field test, in actual classroom situations, such materials, and publish the results of their effectiveness. (69, 70, 71)

Section I

A GENERAL POINT OF VIEW

Evaluation is a common practice in American society. From the worn but sturdy cliché "the unexamined life is not worth living" to the precise timing of the long-distance runner, ours seems to be a culture of assessment, comparison, evaluation. The large issue to which the NEA Task Force on Testing has turned its attention is not so much whether there should be evaluation but what should be its nature, who should conduct it, how should those who conduct evaluation be prepared, and how should the results of evaluation be used.

The Task Force was impressed with the strong thread running throughout its hearings and from the literature of the potential profound effect on human beings' lives of the classifying and labeling characteristics and uses of tests. It was frequently reported that tests are developed and used in ways that serve to keep certain individuals and groups "in their places" near or at the bottom of the social-economic scale and to assure other individuals and groups that they will maintain present high status positions both socially and economically. The Task Force concluded that while its approach to evaluation would be constructive and positive, such destructive characteristics of tests and measurements must be resisted in every way. The use of tests, as Arthur Coombs has prioritized the teaching of reading, must at times be superseded by the development of the students' self-concept.

Because the main charge to the Task Force was to respond to NEA resolutions and new business items on testing and evaluation that appear at the beginning of this report, and particularly to the issues revolving around standardized testing, the Task Force has developed its major efforts, its findings, and its recommendations to those ends.

But the Task Force is aware that the problems of standardized testing are part of a much broader context, are central to the much more complicated fabric of accountability. And woven into that fabric are such other issues as—

1. Performance-based education
2. Performance-based teacher education
3. National and state assessment
4. Evaluation of educational programs and conditions
5. Criteria for teacher certification and recertification
6. Criteria for teacher selection, retention, promotion, and dismissal
7. Other issues in addition to testing that result in displacement and exclusion of students from learning opportunities.

It is the point of view of the Task Force that the united teaching profession must ultimately deal with all of these. But not all can or should be dealt with through the same mechanisms or along identical time lines. For this reason the recommendations for further study are presented in two separate sections:

One dealing with those issues the Task Force believes to be direct testing issues (Section IV);

And a second dealing with other important assessment and decision-making issues, which may need to be dealt with in interlocking NEA programs and projects (Section V).

The Task Force calls attention here to the significance for its work, and for continuing work on testing issues, of the resolutions and items of new business of the 1972 Representative Assembly that address themselves to these issues. The resolutions appear in the front of this report. The Task Force believes that, as stated in Resolution 72-44, the NEA should continue to encourage "the elimination of the use of group standardized intelligence, aptitude, and achievement tests to assess student potential until completion of a critical appraisal, review, and revision of current testing programs." A number of state education associations have already taken action, based on that recommendation, calling for a moratorium on testing in their states.

At the same time, the Task Force is aware that, in some states, statutes mandating testing programs and local school district policies on testing will need to be revised or removed. The Task Force proposes in Section III of this report areas for immediate action by NEA.

Because of the complexity of the tasks that it undertook, the relatively short period of time that it functioned, and the commitment of the NEA to continue to study the testing issues for two more years (1972 Representative Asembly Item of New Business 28), the Task Force emphasized the identification of specific areas for continued in-depth study. The main substance of these areas appears in Section IV.

Section II

THE TASK FORCE BELIEVES . . .

The positions taken below are based on over 30 hours of hearings, survey of the vast literature on testing and evaluation in education, and debate by Task Force members of the issues. While time limitations did not permit exhaustive study or empirical research by the Task Force, the findings are based on expert judgment, experience, and research reported by witnesses representing such groups as teachers, students, minorities, government agencies, college and university personnel, school administrators, testing industry, and a wide variety of professional associations concerned with educational and psychological testing. The Task Force stands on these premises, recognizing, however, that a number of them require further investigation. The nature of such investigation is proposed in sections IV and V.

1. The Task Force believes that some measurement and evaluation in education is necessary.

> A state education association human relations director told the Task Force, "Don't deny testing as an essential area . . . but it must be based on experiences people have had."
>
> Holmen and Docter conclude that ". . . few would argue against allowing schools to give tests to determine what a student has learned in some course of study."[1]
>
> As a representative of a national testing association pointed out, "Descriptions and decisions are going to be made with or without tests. It's inevitable. . . . If we are going to make descriptions and decisions, it makes sense, within limits of costs, to seek the best information."

[1]Holmen, Milton G., and Docter, Richard. Educational and Psychological Testing. New York: Russell Sage Foundation, 1972. p. 13.

2. The Task Force believes that some of the measurement and evaluation tools developed over the years, and currently in use, contain satisfactory validity and reliability requirements and serve useful purposes when properly administered and interpreted.

Teachers reported that individual diagnostic instruments in such basic skill areas as reading and mathematics are helpful in identifying appropriate remedial action. And what is called Item Response Analysis in the Cleveland Public Schools appears to be a promising approach – clusters of item responses are used to develop educational prescriptions in response to identified learning problems. Teachers are treated as the professionals they are in that they are encouraged to select and try alternate teaching resources; that is, they both develop and apply the prescription. A key question asked in the Cleveland plan in analyzing clusters of responses is, "Is this something that should be reasonably attained by the child?"

3. The Task Force believes that certain measurement and evaluation tools are either invalid and unreliable, out-of-date, or unfair and should be withdrawn from use.

The unfairness of some tests to some students was brought to the attention of the Task Force from a variety of sources. A group of minority students told of being placed in special education classes on the basis of being below grade level on standardized achievement tests, placements that could be adjusted only after 3 years. Instances were related of black students' being denied participation in extracurricular activities on the basis of tests. Teachers reported that group tests applied to very small children are unreliable because of the children's varying attention spans and maturity levels.

The Task Force was particularly impressed with substantial testimony to the effect that both standardized achievement and intelligence tests are unfair to bilingual/bicultural students as well as to non-English-speaking and non-standard-English-speaking students. We cite here the following following resolution submitted by the Bay Area Bilingual Education League of California and adopted by the NEA First American and Hispanic Task Force which bears directly on this issue:

RESOLUTION

> Testing of children whose language is other than standard English with instruments that were developed for users of standard English violates the norm and standardization of these instruments and makes the results questionable. We contend that the use of these instruments with children whose language is other than standard English is invalid.
>
> Sufficient evidence now exists to direct us to the development of criterion-referenced assessment systems as a means of improving the accountability of educational programs. These evaluation processes must correspond to local performance objectives.
>
> The development of valid test instruments for bilingual and bicultural children must be directed by qualified bilingual and bicultural personnel in the educational field or in similar fields, to assure that the test instruments will reflect the values and skills of the ethnic and cultural groups being tested.
>
> Whereas currently used standardized tests measure the potential and ability of neither bilingual nor bicultural children and yet are so used and relied upon to count, place and track these children, we resolve that such use of standardized tests be immediately discontinued.

It was also called to the attention of the Task Force that standardized tests discriminate unfairly on the basis of sex.

4. The Task Force believes that the training of those who use measurement and evaluation tools is woefully inadequate and that schools of education, school systems, and the testing industry all must take responsibility for correcting these inadequacies. Such training must develop understanding about the limitations of tests for making predictions about potential learning ability, of their lack of validity in measuring innate characteristics, and their dehumanizing effects on many students. It must also include understanding the students' rights related to testing and the use of test results.

Teachers reported that they are frequently unfamiliar with the tests they are required to administer, the purposes of the overall evaluation programs they are a part of, and the uses that will be made of the results of testing programs. They told the Task Force that neither preservice nor in-service programs for teachers provide adequate preparation for administration and interpretation of tests or prescribing learning activities based on the findings.

Professors of education told the Task Force that the components on tests and measurement in teacher education programs are frequently vague or nearly absent, and that in many institutions there are no requirements for instructions in tests and measurement as a part of teacher education programs. A survey of requirements in the 50 states for instruction in tests and measurements as a prerequisite for teacher licensure showed that only 13 states have such requirements and some of these apply only to specific groups of teacher trainees, e.g., special education and guidance and counseling.

5. The Task Force believes there is overkill in the use of standardized tests and that the intended purposes of testing can be accomplished through less use of standardized tests, through sampling techniques where tests are used, and through a variety of alternatives to tests.

Holmen and Docter[2] estimate that at least 200 million achievement test forms are used each year in the U.S. And this, they report, is only 65 percent of all educational and psychological testing that is carried out. Even though it is difficult to know how much is too much in this arena, it appears to represent three or four standardized tests per student per year. And this is in addition to the millions of teacher-made tests, surveys, inventories, and oral quizzes to which students are subjected annually.

Representatives of the testing industry and others told the Task Force that sampling of student populations could be as effective as the blanket application of tests that is now so common. Some suggested that such procedures, in addition to increasing the assurance of privacy rights, would conserve time, effort, and financial expenditure.

6. The Task Force believes that the National Teacher Examinations are an improper tool and must not be used for teacher certification, recertification, selection, assignment, retention, salary determination, promotion, transfer, tenure, or dismissal.

The Task Force heard testimony that the National Teacher Examinations have been used to license, select, assign, transfer, promote, and dismiss teachers. Research indicates that no single objective tool is highly enough developed for these purposes. It therefore seems apparent that application of the NTE for these purposes represents misuse of the instrument. The Educational Testing Service itself, developer and marketing agent for the Examinations, has acknowledged that some of these purposes constitute misuse of the test.

[2] Ibid., p. 38.

7. The Task Force believes that the results from group standardized tests should not be used as a basis for allocation of federal or state funds.

The Task Force learned that in some states some funds are distributed to schools on the basis of student scores on standardized tests. And some guidelines for proposal development in applying for federal funds require that systemwide testing programs be agreed to as part of eligibility for participation. Since standardized tests apply so unevenly to different groups and individuals and often poorly predict potential learning ability, and since so many of them are incapable of diagnosing the most significant learning difficulties, it would appear that their use for determining which educational programs should be funded and for what students would result in inaccuracy and unfair treatment of some groups and individuals.

8. The Task Force believes that standardized tests should not be used for tracking students.

The issue of tracking in and of itself has been a practice of questionable value for many years. A concentration of studies in the 20's and 30's found little evidence that homogeneous grouping improved student learning. In the 50's, when American schools were being pointed at as contributing to the United States's second position in the space race, tracking was again widely instituted, followed by another concentration of studies on its effects. The findings the second time indicated that in general children who were grouped learned no more than those who were treated heterogeneously. To date no substantial evidence of increased learning as a result of tracking has been produced, yet tracking goes on. Some kinds of special education may be defensible for some students for part of the time on the basis of making teachers' jobs more manageable. But if this or other reasons apply, they should be put forth, rather than that learning is improved. Even then, assignment to special programs should be based on individual student needs determined by individually administered diagnostic instruments, by mutual agreement with parents, and on a part-time and temporary basis. There should be opportunities for students to move back and forth from regular to special programs as their social and emotional needs as well as academic requirements indicate.

9. The Task Force believes that while the purposes and procedures of the National Assessment of Education may have been initially sound, a number of state adaptations of such programs have subverted the original intent and as a result are potentially harmful.

A main purpose of the National Assessment of Educational Progress has been to determine, for representative samples of the American public, levels of understandings and abilities to perform in a variety of areas considered by its developers important for a large majority of the society. The Task Force believes, as reported by an earlier NEA Task Force,[3] "that all Americans need to be educated, and that it is essential to identify the educational needs of our people and to respond to those needs with relevant and effective educational programs, both through formal schooling and through other means." (The NEA Task Force on Compulsory Education, in its report, recommends a number of promising alternatives to present school organization and process for accomplishing the ends.) The Task Force on Testing is supportive of efforts to identify the educational needs of the American people.

But adaptations of national assessment programs in some states are being manifest as statewide testing programs, applied to all students, and used to compare

[3]NEA Task Force on Compulsory Education. Report of the Task Force. Washington, D.C.: NEA, 1973.

population groups, school systems, individual schools, even teachers and students.

Both such applications and the dissemination of the results from them have deleterious effects on students and teachers and evoke inaccurate and negative responses in public understanding of and attitude toward the schools. Members have expressed concern about National Assessment through Resolution 72-13.[4]

10. The Task Force believes that both the content and use of the typical group intelligence test are biased against those who are economically disadvantaged and culturally and linguistically different, and especially against all minority groups.

Hoffman reports, "There is no generally satisfactory method of evaluating human abilities and capabilities, though occasionally it can be done individually with remarkable precision.[5]

Considerable research over the years has led to the conclusion that the most commonly used group intelligence tests measure only one aspect of intelligence—verbal capacity. And even if it were agreed that this aspect is an important predictor of capacity to be successful in the society, conventional intelligence tests still are grossly flawed. For these reasons some have called for complete elimination of group tests of mental ability, including abolishment of the term "IQ."

Scores on tests of mental ability are so influenced by past experience and cultural background that they are highly biased in favor of those groups whose experience and culture the items reflect. The content frequently highly reflects middle-class culture and experience. The tests are often characterized by an ambiguity that confuses those who think critically and in depth. Hoffman[6] reported this more than a decade ago. In addition, the work of Getzels and Jackson, later followed up by Torrance, has shown that intelligence tests reflect mainly the ability to converge on single, predetermined correct answers. An important prerequisite to creativity, the ability to carry on divergent thinking, is not often measured in the typical intelligence test. As Barzun has put it, mechanical tests raise mediocrity above talent.[7]

Edward Casavantes, a prominent Chicano psychologist,[8] told the Task Force that poverty alone is the major factor in causing minority groups to appear to be of less ability than others.

This effect of poverty on IQ is further substantiated by Jane Mercer in a report on her landmark research in which she states that "persons from the lowest socio-economic groups were far more likely to be (considered mentally retarded) than were those from higher status levels."[9]

---

[4]"National Testing and Assessment—72-13." The National Education Association notes that the first report of the National Assessment of Educational Progress on writing, citizenship, and science has been issued.

The Association will continue to resist any attempt to transform assessment results into a national testing program that would seek to measure all students or school systems by a single standard, and thereby impose upon them a single program rather than providing opportunities for multiple programs and objectives. Washington, D.C.: NEA, 1972.

[5]Hoffman, Banesh. "Psychometric Scientism." Phi Delta Kappan, April 1967.

[6]Hoffman, Banesh. The Tyranny of Testing. New York: Crowell-Collier, 1962.

[7]Barzun, Jacques. Teacher in America. Boston: Little, Brown, & Co., 1945. (Doubleday, Anchor Books, 1954.)

[8]Casavantes, Edward, Executive Officer, Association of Psychologists por La Raza, in testimony before the Task Force, March 31, 1973.

[9]Mercer, Jane R. Labeling the Mentally Retarded. Berkeley: University of California Press, 1972.

11. The Task Force believes that the use of the typical intelligence test contributes to what has come to be termed "the self-fulfilling prophecy," whereby students' achievement tends to fulfill the expectations held by others.

The Task Force was impressed by considerable testimony in support of the findings of the Rosenthal and Jacobson study.[10] Where heavy emphasis is placed on intelligence testing, students may tend to be pigeonholed on the basis of tests. Less is expected of those who do less well on the tests. There is little question that teachers' expectations contribute to student performance. Thus, it can be concluded that those who are expected to achieve less actually achieve less, and vice versa.

12. The Task Force believes that test results are too often used by educators, students, and parents in ways that are hurtful to the self-concept of many students.

Holmen and Docter[11] report that of all the criticisms of tests this one is the most difficult to dismiss. Few would deny the importance of a positive self-image to enhance the possibilities for student learning.

13. The Task Force believes that the testing industry must demonstrate significantly increased responsibility for validity, reliability, and up-to-dateness of their tests, for their fair application, and for accurate and just interpretation and use of their results.

The Task Force objects to the strong tendency of representatives of the testing industry to place most of the blame for the problems of testing on test usage and to assume little responsibility for the uses made of their products.

But a prior issue is the responsibility of the industry to ensure relevant content, validity, and reliability in its product. The Task Force was told that some tests remain on the market for many years beyond a time when much of their content has become irrelevant simply because there continues to be a market for them.

Matters of validity and reliability, fair application, and accurate and just interpretation and use are dealt with at other places in this report. It need only be reiterated here that these are joint responsibilities in which the testing industry needs to participate much more than it has in the past.

14. The Task Force believes that the public, and some in the profession, misinterpret the results of tests as they relate to status and needs of groups of students as well as to individual students.

The statistical fact that 50 percent of any population will always end up below the mathematical average ("norm") leads many to believe that being below average means poor quality performance. This is not necessarily so. The mathematical average may or may not be highly related to competent performance. The public, particularly, needs to come to understand that norming processes automatically place half the students below the average, no matter how well they perform. The Task Force heard testimony that the use of Grade Equivalent scores leads to drawing inappropriate conclusions on the part of educators, parents, and students.

15. In summary, the Task Force believes that the major use of tests should be for the improvement of instruction—for diagnosis of learning difficulties and for prescribing

[10] Rosenthal, R., and Jacobson, L. Pygmalion in the Classroom: Teacher Expectations and Student's Intellectual Ability. New York: Holt, 1970.

[11] Holmen and Docter. Ibid., p. 38.

learning activities in response to learning needs. They must not be used in any way that will lead to labeling and classifying of students, for tracking into homogeneous groups as the major determinants to educational programs, to perpetuate an elitism, or to maintain some groups and individuals "in their place" near the bottom of the socioeconomic ladder. In short, tests must not be used in ways that will deny any student full access to equal educational opportunity.

Section III

RECOMMENDATIONS FOR IMMEDIATE ACTION

(1973-74 Year)

1. In the fall of 1973, the NEA should provide to all state affiliates, for communication to all state-affiliated locals, and to agencies and associations concerned with educational testing issues, specific guidelines appropriate for adoption as local school district policy calling for—

   a. Immediate replacement of blanket use of (i.e., application to all students) group standardized achievement tests by sampling when necessary of the various school populations
   b. Provision to local school districts by test suppliers of procedures for using different item samples on different student populations and individuals.

   The Task Force believes that immediate implementation of such procedures will serve the purpose of improving the conditions surrounding rights of privacy of students, and prevent publication of scores conducive to stigmatizing minority and nonminority students. Such procedures would also reduce the inordinate amount of time spent in test administration and scoring.

2. In the fall of 1973, NEA should begin consultation with the National Council on Accreditation of Teacher Education and the American Association of Colleges of Teacher Education to influence revision of the current accreditation standards and school of education curricula to include specific requirements for instruction in tests and measurement for all preservice teacher education programs. In such consultations, topics should include items listed under No. 4 of "The Task Force Believes," p. 40.

3. The NEA should begin consultation with such organizations as NCME, AERA, and APA to consider appropriate revisions to the Standards for Development and Use of Educational and Psychological Tests developed cooperatively to assure the proper development and use of standardized tests.

4. By February 1 of 1974, the NEA should provide to all state affiliates and to agencies and associations concerned with testing issues, for communication to all state-affiliated locals, specific guidelines appropriate for adoption as local school district policy calling for—

   The local development of criterion-referenced tests in all branches of the curriculum as alternatives to current standardized testing programs.

   While the Task Force has been cautioned that the local development of valid and reliable criterion-referenced tests is a complex and time-consuming job, we believe it must be done, and such efforts must get under way immediately.

5. By June 1 of 1974, the NEA should provide to all state affiliates, for communication to all state-affiliated locals, and to agencies and associations concerned with educational testing issues, specific guidelines for minimal content for in-service education programs for teachers and other school staff, including paraprofessionals, on tests and measurement. Such content should include items listed under No. 4 of "The Task Force Believes," p. 40.

6. By June 1974, the NEA should provide to all state affiliates, for communication to all state-affiliated locals, and to other agencies and associations concerned with educational testing issues, news-release type materials for use in educating both educators and the public on the appropriate uses and limitations of test results and familiarization with a range of alternatives to current common testing practices.

Section IV

RECOMMENDATIONS FOR FURTHER STUDY BY THE NATIONAL EDUCATION ASSOCIATION ON TESTING ISSUES

The recommendations that follow are intended to be pursued during the 1973-74 year concurrently with the implementation of those in the preceding section. In addition, the recommendations in this section should be pursued in depth throughout 1974-75, final recommendations for policy and action to be made by the Task Force on Testing to the 1975 NEA Representative Assembly.

Goals for Accomplishment by 1975

The Task Force recommends intensive study leading to specific action recommendations on the following by June 1975:

1. Essential roles and responsibilities of various concerned groups[12] in assuring sound and fair development of evaluation systems

   The term evaluation systems is used here instead of tests because it is urged and expected that a wide variety of alternatives to tests should and can be developed for evaluation purposes. The Task Force was cautioned that alternatives, perhaps even more than conventional tests, must be subjected to rigorous research and test and tryout leading to validation.

2. Essential roles and responsibilities of such groups[12] in assuring appropriate distribution and administration of evaluation systems

3. Essential roles and responsibilities of such groups[12] in assuring accurate and fair interpretation of the results of evaluation systems

4. Essential roles and responsibilities of such groups[12] in assuring relevant and constructive action programs based on the results of the use of evaluation systems.

The above should be considered general goals. Action programs will need to be developed for accomplishing each of the goals. Some programs may be developed that will respond to more than one of the goals.

[12] See page 38, Part I, for listing of groups.

Recommended Areas for In-Depth Study Required To Accomplish the Goals

The categories listed here were identified early in the deliberations of the Task Force and have been refined as the issues were studied and discussed. The Task Force began to have some strong impressions about some of them on which recommendations and actions might be taken. The Task Force speaks out on these in Section II. But as was indicated in Section I, because of the complexity of the subject, the limitation of time, and because, by resolution, the NEA is committed to study the testing issue for two more years the Task Force rather emphasized the identification of areas for in-depth study.

It is recommended that each of the categories below be studied in depth during 1973-74 and that the final recommendations to the 1975 Representative Assembly reflect actions directed to the specific items in each category. The categories are The Student, The Teacher, The Testing Industry, The Government, and Other Agencies and Associations.

I. The Student
- A. Effect of tests on labeling and classifying students in ways that restrict the development of their potential.
- B. Bias in test content that leads to unfair results with some groups on the basis of race, sex, socioeconomic status, bilingual/bicultural, non-English- and non-standard-English-speaking.
- C. Effect of tests on student self-concept.
- D. Effect of tests on the "self-fulfilling prophecy" concept. (See p. 43.)
- E. Degree to which the content and use of tests invades privacy of students.
- F. Degree to which publication of test scores invades the privacy of students.
- G. Degree to which tests affect the more mobile members of the student population.
- H. Degree to which tests contribute to the development in students of limited cognitive styles, e.g., convergent as opposed to divergent thinking. (See p. 42.)
- I. Promise of alternatives for evaluating human capabilities such as the Ertl Index, the Belmont Battery, Test of Logical Thinking.

II. The Teacher
- A. Effects of tests applied to teachers, i.e., professional status, morale, feelings of security.
  1. National Teacher Examinations and other tests applied directly to teachers.
  2. Use of student test results to judge teachers for retention, tenure, promotion. (See p. 40.)
- B. Effect of tests on curriculum development by educators.
- C. Effects of tests on experimentation with and implementation of new ways of teaching.
- D. Effects of teaching to the tests.
- E. Effects of tests on teachers' ability to individualize instruction.
- F. Effects of mandated testing programs on teacher academic freedom.
- G. Effects of use of tests to hold teachers responsible for educational outcomes of students.

III. The Testing Industry
- A. The responsibility of the industry for distribution of valid, reliable, up-to-date products.
- B. The responsibility of the testing industry to withhold tests and services where there is reasonable certainty they will be misused.
- C. The responsibility of the testing industry to provide validation data for specific regions and specific populations.

D. The responsibility of the testing industry to consult with professional organizations in the development of standards of training for test usage and to share in the responsibility for enforcement of the standards.
E. The responsibility of the testing industry to relate testing to curricula and to assure that appropriate methods of evaluation be considered an integral part of curriculum development.
F. Responsibility of the testing industry to conduct in-depth research, test and tryout of its products, and to continuously research their effects throughout the time of their use and to share information with the profession and the public on the extent of this research effort.

IV. The Government
A. The responsibility of government at all levels (national, state, and local, including local school boards) to assure that biased evaluation systems, and particularly the results of standardized tests, are not used for the allocation of funds.
B. The responsibility of government at all levels to assure that the results of national and state assessment programs are not used for labeling and classifying students or for judging teachers.
C. The responsibility of government at all levels to assure that national and state assessment programs do not lead to national and state curricula.
D. The responsibility of government at all levels to assure that the results of tests are not publicized in ways that violate the privacy of individuals or stigmatize specific populations, school building units, or school attendance areas.
E. The responsibility of government at all levels to assure that the results of tests are not used in any way to promote segregation among or within schools, or to negatively affect teacher assignment.
F. The responsibility of government at national and state levels to provide standards of licensure for test developers and producers.

V. Other Organizations and Associations
A. The responsibility of national professional associations and other organizations associated with testing to fully involve, in a formal relationship, the organized teaching profession in all activities leading to the development of all policy, guidelines, and procedures related to test development and usage.
B. The activities of the College Entrance Examination Board in influencing college admission policy through the use of tests, including the effects of the work of its Commission on Tests.
C. Colleges' and universities' responsibilities in developing and implementing alternatives to present testing arrangements for admission to higher education.

1. In this regard, the strengths of present open admissions programs should be studied and recommendations made on the basis of findings on their success and promise.
2. Examination of the effects of the College Means Admission Program on students and the institutions.

D. The responsibility of the Education Commission of the States to assure that appropriate guidelines and cautions accompany the dissemination of both the instruments for and results from the National Assessment of Educational Progress.
E. Further cooperation with such testing reform efforts as The National Project on Testing in Education and The National Project on Educational Testing.

Some Recommended Actions for Accomplishing the Goals

The recommendations that follow represent some, not all, specific actions to be taken that will contribute to accomplishing one or more of the four goals stated previously (p. 45). These actions will need to become part of broader programs. It is expected that the continuing work of the Task Force on Testing will give high priority to spelling out such programs. (The numbers in brackets following the items indicate the goal or goals which the particular action will contribute to accomplishing):

1. Develop model standards of training and experience for state certification requirements for all those who administer and/or use test results in the school. (2)

2. Develop action plans to assure better control of test development and distribution by the testing industry through—
   a. Influencing appropriate federal and state agencies to better protect test consumers.
   b. Specifically, reducing legal barriers, including restrictions on RTC's refusing test sales to unqualified users. (2)
   c. Support legal action where appropriate to challenge misuse of tests and violation of rights of educators and students.

3. Develop a program for broad publicizing of guidelines for collection, maintenance, and dissemination of pupil records, including those recommended in—
   a. NEA Code of Student Rights and Responsibilities.
   b. Guidelines for the Collection, Maintenance and Dissemination of Pupil Records, a report of the Russell Sage Foundation. (3)

4. Extend the guidelines cited in #3 above by developing model policy statements on the publication of and general dissemination of test scores. (3)

5. Recommendation of specific alternatives to standardized tests appropriate to the evaluation of students and educators.

Section V

RECOMMENDATIONS FOR FURTHER STUDY BY THE NATIONAL EDUCATION ASSOCIATION ON OTHER ASSESSMENT-RELATED ISSUES

As was pointed out in an earlier section, testing is a part of a much broader fabric that has come to be called accountability. Accountability means different things depending on who is defining it. But to many in the public and some in the profession it has to do directly with producing specific outcomes with students, particularly in such basic skill areas as reading and mathematics. This aspect of accountability is obviously directly related to testing in that student performance is most often measured by the use of tests, particularly standardized tests. Other test-related issues that also are important to the accountability movement include—

1. National and state assessment programs
2. Performance-based education
3. Performance-based teacher education

4. Management by objectives
5. Program, planning, budgeting, evaluation systems
6. Evaluation of educational programs and conditions
7. Criteria for teacher certification and recertification
8. Criteria for teacher selection, retention, salary determination, promotion, and dismissal.

Each of these is in some way related to the other and to evaluation, and to tests and measurement. But the Task Force believes that several of these may not fall directly within the purview of the Task Force on Testing.

We recommend that, as the testing issues continue to be studied and acted upon (as recommended in the preceding section), #1 above, the issues surrounding national and state assessment continue to be considered in addition to other testing issues.

The others in the above list should be dealt with as follows:

1. Numbers 2, 3, and 7 are of extreme importance to the teaching profession and should become the concern of a national task force appointed by the NEA president, with an appropriate secretariat and with its work coordinated with the NEA program budget.

> The Testing Task Force learned of three national efforts on performance-based education and teacher education. None of these, at present, has had substantial input from the organized teaching profession. One of them, spearheaded by the Educational Testing Service, threatens to become a major effort to centralize coordination of the entire performance-based movement.

The Task Force strongly urges that appropriate administrative assignments be made as soon as possible so that staff can begin working toward resolving those test-related issues which do not fall under the direct charge of the Task Force on Testing. In addition, all test-related issues should be vigorously pursued as directed by the appropriate resolutions and items of new business dealing with testing.

RESOURCE MATERIALS

Report of the Commission on Tests, I. Righting the Balance, College Entrance Examination Board, New York, 1970.

Report of the Commission on Tests, II. Briefs, College Entrance Examination Board, New York, 1970.

An Investigation of Sources of Bias in the Prediction of Job Performance, A Six-Year Study —Proceedings of Invitational Conference, The Barclay Hotel, New York, June 22, 1972, Educational Testing Service, Princeton, New Jersey.

Ethnic Isolation of Mexican Americans in the Public Schools of the Southwest, Report I, U.S. Commission on Civil Rights, Mexican American Education Study, April 1971.

The Unfinished Education: Outcomes for Minorities in the Five Southwestern States, Report II, Mexican American Educational Series, A Report of the U.S. Commission on Civil Rights, October 1971.

The Excluded Student, Educational Practices Affecting Mexican Americans in the Southwest, Report III, U.S. Commission on Civil Rights, May 1972.

Mexican American Education in Texas: A Function of Wealth, Report IV, Mexican American Education Study, U.S. Commission on Civil Rights, August 1972.

Methodological Appendix of Research Methods Employed in the Mexican American Education Study, U.S. Commission on Civil Rights, January 1972.
Evaluation in the Inner City, Report of an Invitational Conference on Measurement in Education, April 24-25, 1969, Philadelphia, Pennsylvania.
Metropolitan Achievement Tests Special Report, 1970 Edition, Selected Items Revised or Eliminated from the Metropolitan as a Result of Possible Bias Against Minority Groups, Report No. 23 issued by Test Department, Harcourt Brace Jovanovich, Inc., September 1972.
The Use of Standardized Instruments with Urban and Minority-Group Pupils, Thomas J. Fitzgibbon, Test Department, Harcourt Brace Jovanovich, Inc.
Standardized Test Report, 1973 Delegate Assembly, Iowa State Education Association—Informational Services Division, December 1972.
Standards for Development and Use of Educational and Psychological Tests, Third Draft (formerly called Standards for Educational and Psychological Tests and Manuals).
The Responsible Use of Tests: A Position Paper of AMEG, APGA, and NCME, reproduced by permission from Measurement and Evaluation in Guidance, Vol. 5, No. 2, July 1972.
Project Opportunity, A Demonstration Guidance Project for Minority Poverty Youth, March 1973, College Entrance Examination Board.
Pluralistic Diagnosis in the Evaluation of Black and Chicano Children: A Procedure for Taking Sociocultural Variables into Account in Clinical Assessment, presented at the Meetings of the American Psychological Association, Washington, D.C., September 3-7, 1971, by Jane R. Mercer, Associate Professor, Sociology, University of California.
On the Explanation of Racial-Ethnic Group Differences in Achievement Test Scores, George W. Mayeske, U.S. Office of Education, Washington, D.C.
School Psychology and the Mexican American, Dr. Steve G. Moreno, AMAE Newsletter, May 1972.
Position of American Association of School Administrators, George B. Redfern.
Educational Measurement of What Characteristic of Whom (or What) by Whom and Why, Jack C. Merwin, University of Minnesota—Journal of Educational Measurement, Volume 10, No. 1, Spring 1973.
A Moratorium on Public School Testing, The Position of a Professor of Measurement and Evaluation, Andrew L. King, D.C. Teachers College.
Studies Relative to "Improvement" in the Cognitive Performance of Minority Children Under Special Conditions, Edward J. Casavantes.
A Litany of Laments: Elementary and Secondary School Factors Affecting Minority Group Resources for the Health and Mental Health Professions, Edward J. Casavantes (paper presented at a symposium at the American Association for the Advancement of Science Convention on Minority Group Resources for the Health Professions, December 27, 1972).
Summer "Experience in Publishing" Training Fellowships, Harcourt Brace Jovanovich Test Department Memorandum, March 15, 1973.
Testimony – Donald Ross Green, Director of Research, CTB/McGraw-Hill, Monterey, California.
Testimony – James H. Ricks, Jr., The Psychological Corporation.
Testimony – Raphael Minsky, Supervisor of Psychological Services, Montgomery County Public Schools, Rockville, Maryland.
Testimony – Winton H. Manning, Educational Testing Service.
Testimony – Lois J. Wilson, New York State United Teachers Assistant Executive Secretary for Studies and Professional Services and Chairperson, NEA Human Relations Council to TTF.
Testimony – Yvonne Burkholz, Dade County Classroom Teachers' Association, Inc.
Brief Report on the Teaching Situation in Alabama, Anthony Butler, Past President, Alabama State Teachers Association.

Points on Testing Requiring Consideration by TTF, Miss Carol Wick, Varying Exceptionalities Resource Teacher, Nashville, Tennessee.

Summary of Statement by Dr. Robbins Barstow, Director of Professional Development, Connecticut Education Association.

Teachers Association of Anne Arundel County, Inc. Action Report, Volume 5, No. 3, December 1972.

Statement on Standardized Tests by Richard C. Gordon, President, Virginia Education Association.

Misc. News Stories on Release of Test Scores: "Poor Marks for the City," New York Times, 3/25/73; "Reading Scores Decline in the City Schools Again"; "Schools in Fairfax Boast of Scores," Washington Post; "Below-Grade Publicity About Reading Scores," NEW York Times, 3/25/73.

Informal Report on Crownpoint..., Eastern Navajo: Eleven Schools' Reactions to the Testing Program.

Reading and Testing: "One Cause of Reading Failure Is Reading Failure," New York Teacher (Magazine Section), an original article by Deborah Meier.

Teacher Education and Professional Standards, Commission of Washington Education Association, Vol. 1, No. 1, February 1973.

The Scheelhaase Case: Scheelhaase vs. Woodbury Central Community School District, et al.

Molville Record article published March 4, 1971, "Individual Basic Skills Scores Reported."

Educational and Psychological Testing—A Study of the Industry and Its Practices, Milton G. Holmen and Rochard F. Docter, Russell Sage Foundation, New York, 1972.

Journal of the Association of Mexican-American Educators, Inc., Vol. 1, No. 1, May 1973.

Violations of Human and Civil Rights: Tests and Use of Tests—Report of the Tenth National Conference on Civil and Human Rights in Education, Washington, D.C. 1972.

Persons Who Testified Before the Task Force, February 17-18, 1973

Doris L. Allan
700 East Broad Street
Columbus, Ohio 43215

Dr. Robbins Barstow
21 Oak Street
Hartford, Connecticut 06106

Frank E. Burdick
Wisconsin Education Association
222 West Washington Avenue
Marison, Wisconsin 53703

Yvonne Burkholz
1809 Brickell Avenue
Miami, Florida 33129

Gloria J. Busch
94-D Fairway Drive
Wethersford, Connecticut 06109

Anthony S. Butler
P.O. Box 177
Lisman, Alabama 36912

C. Sullivan Culver
4513 N.W. 33rd Avenue
Miami, Florida 33142

Warren Cummings
New Jersey Education Association
2 Barrett Avenue
Newton, New Jersey 07860

Walter Eilers
Teacher's Association of Anne Arundel County
6 Dock Street
Annapolis, Maryland 21401

Catherine Garay
2070 MacArthur Street
San Pedro, California 90732

Reynaldo P. Garay, President
Harbor Chapter
Association of Mexican-American Educators, Inc.
2070 MacArthur Street
San Pedro, California 90732

Richard C. Gordon, President
Virginia Education Association
116 Third Street South
Richmond, Virginia 23219

Robert T. Greene
Special Assistant to State Superintendent
State Department of Education
Richmond, Virginia 23216

J. Shelby Guss
Regional Director
Virginia Education Association
116 South Third Street
Richmond, Virginia 23219

Mrs. Mae Sue Henry
North Carolina Association of Classroom Teachers
1204-C Schant Drive
Raleigh, North Carolina 27606

J. Jackson
New Jersey Education Association
25 Wayne Avenue
East Orange, New Jersey 07018

John Lozano
389 - 61st
Oakland, California 94618

Dennis Manzanares
Associate Vice-President, Student NEA
1203 Barcelona Lane
Santa Fe, New Mexico 87501

Charles McLeod, Minority Counselor
Virginia Commonwealth University
920 West Franklin Street
Richmond, Virginia 23220

William M. Mondale, President-Elect
Minnesota Education Association
4030 Jay Lane
White Bear Lake, Minnesota 55110

E. B. Palmer
P.O. Box 27347
Raleigh, North Carolina 27611

Mrs. Edna C. Richards
Acting Director
North Carolina Association of Classroom Teachers
111 West Morgan Street
Raleigh, North Carolina 27610

Conrad A. Romero
Colorado Department of Education
State Office Building
Denver, Colorado 80203

Mary Elizabeth Sievert
President
Davenport Education Association
2707 East Hayes
Davenport, Iowa 52803